Memory technology evolution: an overview of system memory technologies

technology brief, 8th edition

Abstract 2

Introduction 2

Basic DRAM operation 2
- DRAM storage density and power consumption 4
- Memory access time 5
- System bus timing 5
- Memory bus speed 6
- Burst mode access 6

SDRAM technology 7
- Bank interleaving 7
- Increased bandwidth 7
- Registered SDRAM modules 8
- DIMM configurations 8
  - Single-sided and double-sided DIMMs 8
  - Single-rank, dual-rank, and quad-rank DIMMs 8
- DIMM error detection/correction technologies 10
  - The increasing possibility of memory errors 10
  - Basic ECC memory 11
  - Advanced ECC memory 12
- Memory protection technologies 13
  - Online spare memory mode 13
  - Mirrored memory mode 13
  - Lockstep memory mode 14
  - Memory protection mode summary 14

Advanced memory technologies 15
- Double data rate SDRAM technologies 15
  - DDR-1 15
  - DDR-2 17
  - DDR-3 18
  - Module naming convention and peak bandwidth 18
- Fully-buffered DIMMs 19
- Rambus DRAM 21

Importance of using HP-certified memory modules in ProLiant servers 23

Conclusion 23

For more information 24

Call to action 24

Downloaded from www.Manualslib.com manuals search engine

Abstract

Both the widening performance gap between processors and memory and the growth of memory-intensive business applications are driving the need for better memory technologies for servers and workstations. Consequently, several memory technologies are on the market at any given time. HP evaluates developing memory technologies in terms of price, performance, and backward compatibility, and implements the most promising technologies in ProLiant servers.

This paper summarizes the evolution of memory technology and provides an overview of some the newest memory technologies that HP is evaluating for servers and workstations. The purpose is to allay some of the confusion about the performance and benefits of the dynamic random access memory (DRAM) technologies on the market.

Introduction

Processors use system memory to temporarily store the operating system, applications, and data they use and manipulate. Therefore, application performance and data reliability are intrinsically tied to the speed and bandwidth of the system memory. Over the years, these factors have driven the evolution of system memory from asynchronous DRAM technologies, such as Fast Page Mode (FPM) memory and Extended Data Out (EDO) memory, to high-bandwidth synchronous DRAM (SDRAM) technologies. Yet, system memory bandwidth has not kept pace with improvements in processor performance, thus creating a "performance gap." If processor and memory performance continue to increase at current rates, the performance gap between them will widen.

The processor-memory performance gap is important because the processor remains idle while it waits for data from system memory. This performance gap prevents many applications from effectively using the full computing power of modern processors. In an attempt to narrow the performance gap, the industry vigorously pursues the development of new memory technologies. HP works with Joint Electronic Device Engineering Council (JEDEC) memory vendors and chipset developers during memory technology development to ensure that new memory products fulfill customer needs for reliability, cost, and backward compatibility.

This paper describes differences in price, performance, and compatibility of DRAM technologies. Some descriptions are very technical. For readers who are not familiar with memory technology, the paper begins with a description of basic DRAM operation and terminology.

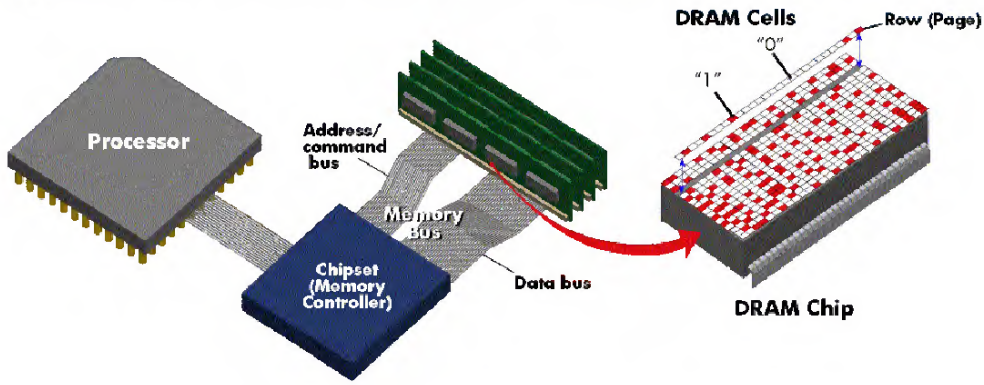
Basic DRAM operation

Before a computer can perform any useful task, it copies applications and data from the hard disk drive to the system memory. Computers use two types of system memory—cache memory and main memory. Cache memory consists of very fast static RAM (SRAM) and is usually integrated with the processor. Main memory consists of DRAM chips on dual inline memory modules (DIMMs) that can be packaged in various ways depending on system form factor.

Downloaded from www.Manualslib.com manuals search engine

Each DRAM chip cantains millians af memary lacatians (ar cells), which are arranged in a matrix af raws and calumns (Figure 1). Peripheral circuitry an the DIMM reads, amplifies, and transfers the data fram the memary cells ta the memary bus. Each DRAM raw, called a page, cansists af several DRAM cells. Each DRAM cell an a page cantains a capacitar capable af staring an electrical charge far a very shart time. A charged cell represents a "1" data bit, and an uncharged cell represents a "0" data bit. Ta prevent the capacitars fram discharging, they must be recharged (refreshed) thausands af times per secand ta maintain the validity af the data. Refresh mechanisms are described later in this sectian.

Figure 1. Representatian af a single DRAM chip an a DIMM



NOTE: Some CPU designs have the memory controller function integroted into processor module

The memary subsystem aperates at the memary bus speed. Typically, ta access a DRAM cell, the memary cantraller sends electranic address signals that specify the raw address and calumn address af the target cell. The memary cantraller sends these signals ta the DRAM chip by way af the memary bus. The memary bus cansists af twa sub-buses: the address/cammand bus and the data bus.

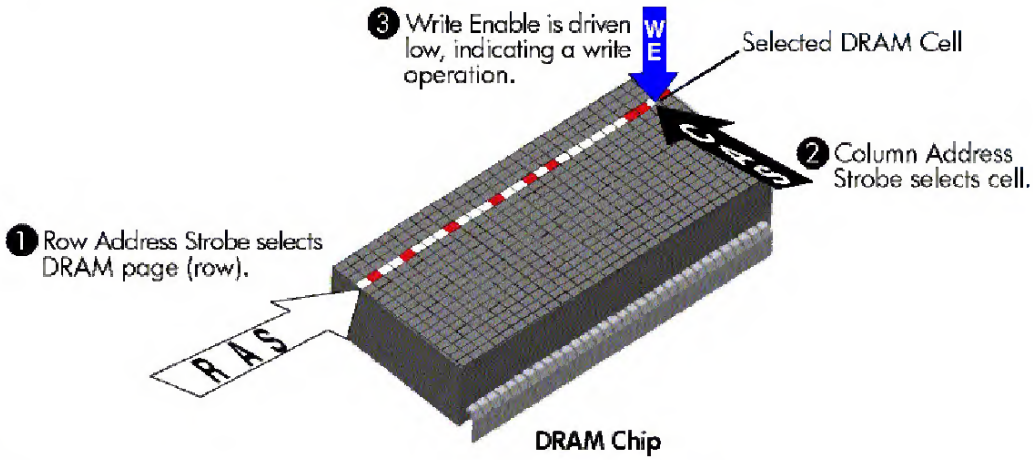
The data bus is a set af lines (traces) that carry the data ta and fram DRAM. Each trace carries ane data bit at a time. The thraughput (bandwidth) af the data bus depends an its width (in bits) and its frequency. The data width af a memary bus is usually 64-bits, which means that the bus has 64 traces, each af which transparts ane bit at a time. Each 64-bit unit af data is called a data ward.

The address partian af the address/cammand bus is a set af traces that carry signals identifying the lacatian af data in memary. The cammand partian af the address/cammand bus canveys instructians such as read, write, ar refresh.

Downloaded from www.Manualslib.com manuals search engine

When FPM or EDO memory *writes* data to a particular cell, the memory controller selects the location in which to write the data. The memory controller first selects the page by strobing the Row Address onto the address/command bus. It then selects the exact location by strobing the Column Address onto the address/command bus (see Figure 2). These actions are called Row Address Strobe (RAS) and Column Address Strobe (CAS). The Write Enable (WE) signal is activated at the same time as the CAS to specify that a write operation is to be performed. The memory controller then drives the data onto the memory bus. The DRAM devices capture (latch) the data and store it into the respective cells.

Figure 2. Representation of a write operation for FPM or EDO RAM


3 Write Enable is driven low, indicating a write operation.
WE
Selected DRAM Cell
2 Column Address Strobe selects cell.
CAS
1 Row Address Strobe selects DRAM page (row).
RAS
DRAM Chip

During a DRAM read operation, RAS followed by CAS are driven onto the memory bus. The WE signal is held inactive, indicating a read operation. After a delay called CAS Latency, the DRAM devices drive the data onto the memory bus.

DRAM cannot be accessed during a refresh. If the processor makes a data request during a DRAM refresh, the data will not be available until the refresh completes. There are many mechanisms to refresh DRAM, including RAS only refresh, CAS before RAS (CBR) refresh, and Hidden refresh. CBR, which involves driving CAS active before driving RAS active, is used most often.

DRAM storage density and power consumption

The storage capacity (density) of DRAM is inversely proportional to the cell geometry. In other words, storage density increases as cell geometry shrinks. Over the past few years, improvements in DRAM storage density have increased capacity from almost 1 kilobit (Kb) per chip to 2 gigabit (Gb) per chip. In the near future, it is expected that capacity will increase even further to 4 Gb per chip.

The industry-standard operating voltage for computer memory components was originally 5 volts. However, as cell geometries decreased, memory circuitry became smaller and more sensitive. Likewise, the industry-standard operating voltage decreased. Today, computer memory components operate at 1.8 volts, which allows them to run faster and consume less power.

Downloaded from www.Manualslib.com manuals search engine

Memory access time

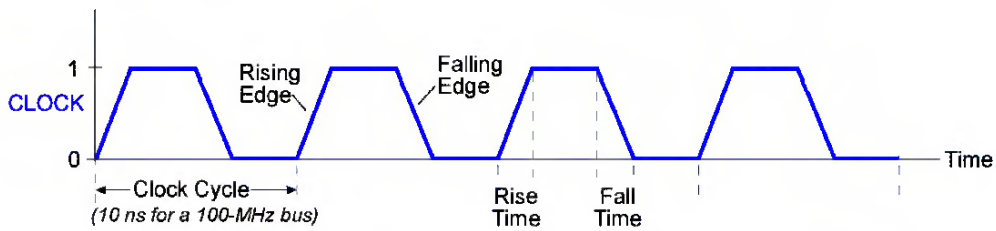
The length of time it takes for DRAM to produce the data, from the CAS signal until the data is available on the data bus, is called the memory access time or CAS Latency. For asynchronous DRAM, memory access time is measured in billionths of a second (nanoseconds, ns). For synchronous DRAM, the memory access time is measured by the number of memory bus clocks.

System bus timing

A system bus clock controls all computer components that execute instructions or transfer data. Typically, the system chipset controls the speed, or frequency, of the system bus clock and thus regulates the traffic between the processor, main memory, PCI bus, and other peripheral buses.

The bus clock is an electronic signal that alternates between two voltages (designated as "0" and "1" in Figure 3) at a specific frequency. The bus frequency is measured in millions of cycles per second, or megahertz (MHz). During each clock cycle, the voltage signal transitions from "0" to "1" and back to "0." A complete clock cycle is measured from one rising edge to the next rising edge. Data transfer along the memory bus can be triggered on either the rising edge or falling edge of the clock signal.

Figure 3. Representation of a bus clock signal



The clock cycle time is measured in nanoseconds (ns, 1 x 10^{-9} second) and the clock frequency is measured in megahertz (MHz, millions of cycles per second). The clock frequency is the reciprocal of the clock cycle time. For example, if the clock cycle is 10 ns then the clock frequency is 1/(0.000000010)=100 MHz.

Over time, some computer components have gained in speed more than others have. For this reason, the components in a typical server are controlled by different clocks that run at different, but related, speeds. These clocks are created by using various clock multiplier and divider circuits to generate multiple signals based on the main system bus clock. For example, if the main system bus operates at 100 MHz, a divider circuit can generate a PCI bus frequency of 33 MHz (system clock ÷ 3) and a multiplier circuit can generate a processor frequency of 400 MHz (system clock x 4). Computer components that operate in whole multiples of the system clock are termed synchronous because they are "in sync" with the system clock.

Synchronous components operate more efficiently than components that are not synchronized (asynchronous) with the system bus clock. With asynchronous components, either the rest of the system or the component itself must wait one or more additional clock cycles for data or instructions due to clock resynchronization. In contrast, synchronized components know on which clock cycle data will be available, thus eliminating these timing delays.

Downloaded from www.Manualslib.com manuals search engine

Memory bus speed

The speed of the DRAM is not the same as the true speed (or frequency) of the overall memory subsystem. The memory subsystem operates at the memory bus speed, which may not be the same frequency (in MHz) as the main system bus clock. The two main factors that control the speed of the memory subsystem are the memory timing and the maximum DRAM speed. The server BIOS sets the memory system speed based on the fastest common frequency of the loaded DIMMs.

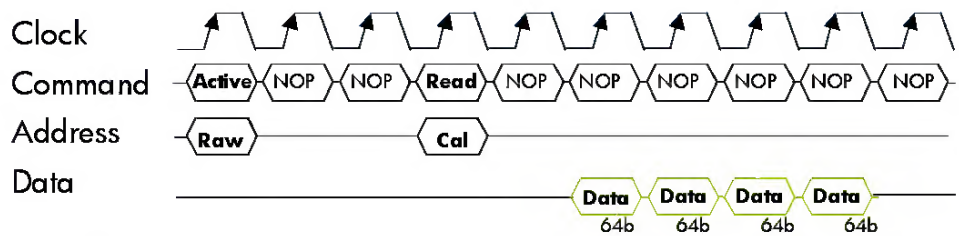
Burst mode occess

The original DRAM took approximately six system-bus clock cycles for each memory access. During memory access, first the RAS and CAS and then 64 bits of data were transferred through the memory bus. The next sequential address access required a repeat of the RAS-CAS-Data sequence. As a result, most of the overhead occurred while transferring row and column addresses, rather than the data.

FPM and EDO improved performance by automatically retrieving data from sequential memory locations on the assumption that they would also be requested. Using this process called burst mode access, four consecutive 64-bit sections of memory are accessed, one after the other, based on the address of the first section. So instead of taking six clock cycles to access each of the last three 64-bit sections, it may take from one to three clock cycles each (see Figure 4).

Burst mode access timing is normally stated in the format "x-y-y-y" where "x" represents the number of clock cycles to read/write the first 64 bits and "y" represents the number of clock cycles required for the second, third, and fourth reads/writes. For example, prior to burst mode access, DRAM took up to 24 clock cycles (6-6-6-6) to access four 64-bit memory sections. With burst mode access, three additional data sections are accessed with every clock cycle after the first access (6-1-1-1) before the memory controller has to send another CAS.

Figure 4. Burst mode access



NOTE: NOP = No Operation Instruction

Downloaded from www.Manualslib.com manuals search engine

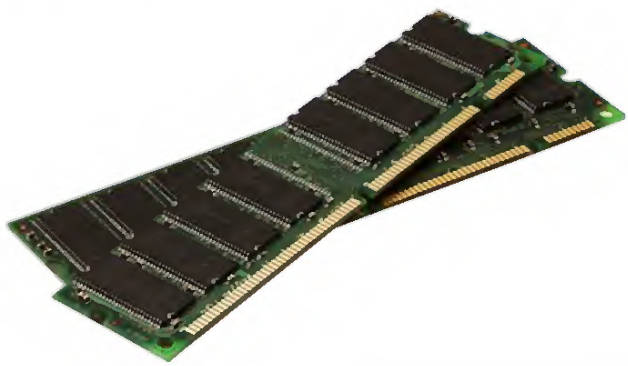
SDRAM technology

FPM and EDO DRAMs are cantralled asynchranausly, that is, withaut a memary bus clack. The memary cantraller determines when ta assert signals and when ta expect data based an absalute timing. The inefficiencies af transferring data between a synchranaus system bus and an asynchranaus memary bus results in langer latency.

Cansequently, JEDEC—the electranics industry standards agency far memary devices and madules—develaped the synchranaus DRAM standard ta reduce the number af system clack cycles required ta read ar write data. SDRAM uses a memary bus clack ta synchranize the input and autput signals an the memary chip. This simplifies the memary cantraller and reduces the latency fram CPU ta memary.

In additian ta synchranaus aperatian and burst made access, SDRAM has ather features that accelerate data retrieval and increase memary capacity: multiple memary banks, greater bandwidth, and register lagic chips. Figure 5 shaws SDRAM DIMMs with twa key natches that prevent incarrect installatian and indicate the type af DRAM madule.

Figure 5. SDRAM DIMM with twa natches



Bank interleaving

SDRAM divides memary inta twa ta faur banks far simultaneaus access ta mare data. This divisian and simultaneaus access is knawn as interleaving. Using a natebaak analagy, two-way interleaving is like dividing each page in a natebaak inta twa parts and having twa assistants ta each retrieve a different part af the page. Even thaugh each assistant must take a break (be refreshed), breaks are staggered sa that at least ane assistant is warking at all times. Therefare, they retrieve the data much faster than a single assistant cauld get the same data fram ane whale page, especially since na data can be accessed when a single assistant takes a break. In ather wards, while ane memary bank is being accessed, the ather bank remains ready ta be accessed. This allaws the pracessar ta initiate a new memary access befare the previaus access campletes and results in cantinuaus data flaw.

Increased bandwidth

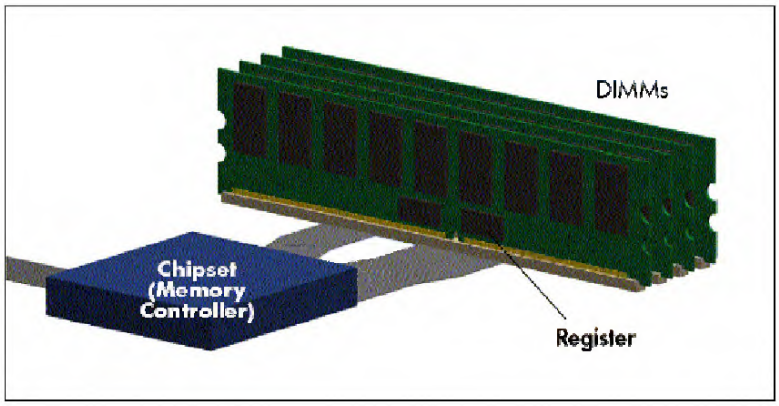
The bandwidth capacity af the memary bus increases with its width (in bits) and its frequency (in MHz). By transferring B bytes (64 bits) at a time and running at 100 MHz, SDRAM increases memary bandwidth ta B00 MB/s, 50 percent mare than EDO DRAMs (533 MB/s at 66 MHz).

Downloaded from www.Manualslib.com manuals search engine

Registered SDRAM modules

To achieve higher memory subsystem capacity, some DIMMs have register logic chips (registers) that act as a pass-through buffer for address and command signals (Figure 6). Registers prevent the memory controller from having to drive the entire arrangement of DRAM chips on each module. Rather, the memory controller drives only the loading of the registers on each module. The register on each DIMM re-drives the address and command signals to the appropriate DRAM chip. Simultaneously, a phase lock loop chip on the registered DIMM generates a second clock signal that runs synchronously with the system bus clock and eliminates the need for the system bus clock signal from having to drive all the DRAM chips. It also allows adding more memory modules to the memory bus to increase memory capacity.

Figure 6. Registered DIMMs



NOTE: Some CPU designs include the memory controller function integroted into processor module.

DIMM configurations

Single-sided and double-sided DIMMs

Each DRAM chip on a DIMM provides either 4 bits or 8 bits of a 64-bit data word. Chips that provide 4 bits are called x4 (by 4), and chips that provide 8 bits are called x8 (by 8). It takes eight x8 chips or sixteen x4 chips to make a 64-bit word, so at least eight chips are located on one or both sides of a DIMM. However, a standard DIMM has enough room to hold a ninth chip on each side. The ninth chip is used to store 4 bits or 8 bits of Error Correction Code (ECC).

An ECC DIMM with all nine DRAM chips on one side is called single-sided, and an ECC DIMM with nine DRAM chips on each side is called double-sided (Figure 7). A single-sided x8 ECC DIMM and a double-sided x4 ECC DIMM each create a single block of 72 bits (64 bits plus 8 ECC bits). In both cases, a single chip-select signal from the memory controller activates all the chips on the DIMM. In contrast, a double-sided x8 DIMM (bottom illustration) requires two chip-select signals to access two 72-bit blocks on two sets of DRAM chips.

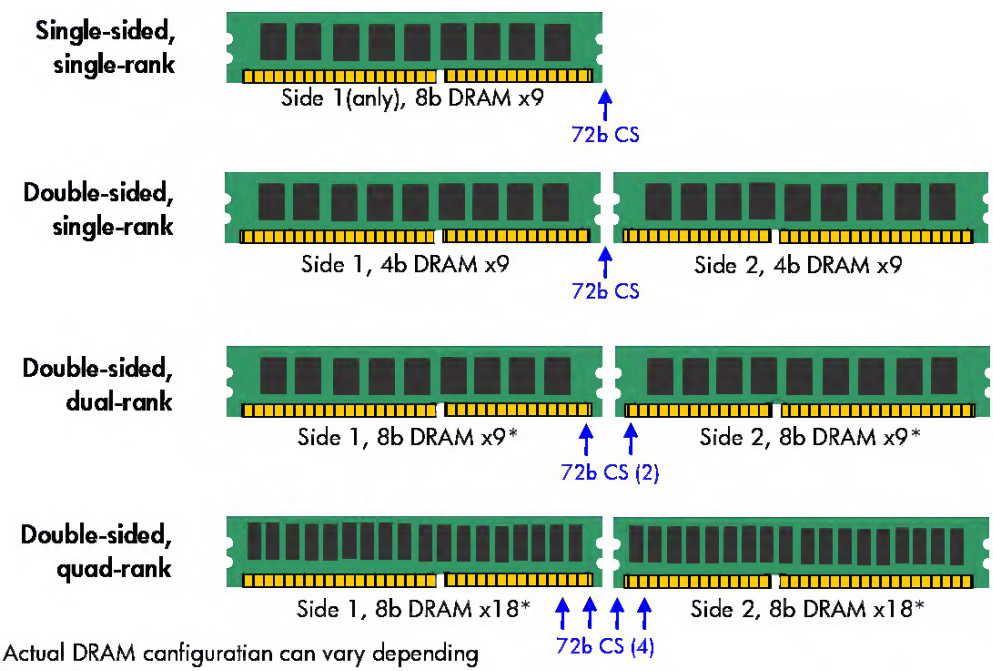
Single-rank, dual-rank, and quad-rank DIMMs

In addition to single-sided and double-sided configurations, DIMMs are classified as to rank. A memory rank is defined as an area or block of 64-bits (72 bits for ECC memory) created by using some or all of the DRAM chips on a DIMM.

Downloaded from www.Manualslib.com manuals search engine

A single-rank ECC DIMM (x4 or x8) uses all of its DRAM chips to create a single block of 72 bits, and all the chips are activated by one chip-select (CS) signal from the memory controller (top two illustrations in Figure 7). A dual-rank ECC DIMM produces two 72-bit blocks from two sets of DRAM chips on the DIMM, requiring two chip-select signals. The chip-select signals are staggered so that both sets of DRAM chips do not contend for the memory bus at the same time. Quad-rank DIMMs with ECC produce four 72-bit blocks from four sets of DRAM chips on the DIMM, requiring four chip-select signals. Like the dual-rank DIMMs, the memory controller staggers the chip-select signals to prevent the four sets of DRAM chips from contending for the memory bus at the same time.

Figure 7. Single-sided and double-sided DDR SDRAM DIMMs and corresponding DIMM rank



* Actual DRAM configuration can vary depending on DRAM density and configuration

Memory ranks are not new, but their role has become more critical with the advent of new chipset and memory technologies and growing server memory capacities. Dual-rank DIMMs improve memory density by placing the components of two single-rank DIMMs in the space of one module. The chipset considers each rank as an electrical load on the memory bus. At slower bus speeds, the number of loads does not adversely affect bus signal integrity. However, for faster memory technologies such as DDR2-667, there is a maximum number of ranks that the chipset can drive. For example, if a memory bus on a server has four DIMM slots, the chipset may only be capable of supporting two dual-rank DIMMs or four single rank DIMMs. If two dual-rank DIMMs are installed, then the last two slots must not be populated. To compensate for the reduction in the number of DIMM slots on a bus at higher speeds, modern chipsets employ multiple memory buses.

If the total number of ranks in the populated DIMM slots exceeds the maximum number of loads the chipset can support, the server may not boot properly or it may not operate reliably. Some systems check the memory configuration while booting to detect invalid memory bus loading. When an invalid memory configuration is detected, the system stops the boot process to avoid unreliable operation.

Downloaded from www.Manualslib.com manuals search engine

To prevent this and other memory-related problems, HP urges customers to use only HP-certified DIMMs, which are available in the memory option kits for each ProLiant server (see the "Importance of using HP-certified memory modules in ProLiant servers" section).

Another important difference between single-rank and dual-rank DIMMs is cost. Typically, memory costs increase with DRAM density. For example, the cost of an advanced, high-density DRAM chip is typically more than twice that of a conventional DRAM chip. Because large capacity single-rank DIMMs are manufactured with higher-density DRAM chips, they typically cost more than dual-rank DIMMs of comparable capacity.

DIMM error detection/correction technologies

Memory modules used in servers are inherently susceptible to memory errors. As described earlier, each DRAM chip stores data in columns and rows of capacitors (memory cells) that must be continuously recharged (refreshed) to preserve the data. The operating voltage of the memory device determines the level of the electrical charge. However, if a capacitor's charge is affected by some external event, the data may become incorrect. Such memory errors can cause applications and operating systems to crash and can result in the permanent loss of business data.

Memory errors are classified by the number of bits that are affected—single-bit or multi-bit—and the cause of error. A 64-bit wide data bus transports 64 bits at a time. These 64 bits constitute an ECC data word. An error in one bit of a data word is a single-bit error. An error in more than one bit of a data word is a multi-bit error.

Depending on the cause, a memory error is referred to as either a hard or soft error. A hard error is caused by a broken or defective piece of hardware, so the device consistently returns incorrect results. For example, a memory cell may be stuck so that it always returns "0" bit, even when a "1" bit is written to it. Hard errors can be caused by DRAM defects, bad solder joints, connector issues, and other physical issues. Soft errors are more prevalent. They occur randomly when an electrical disturbance near a memory cell alters the charge on the capacitor. A soft error does not indicate a problem with a memory device because once the stored data is corrected (for example, by a write to a memory cell), the same error does not recur.

The increasing possibility of memory errors

Two trends increase the likelihood of memory errors in servers: expanding memory capacity and increasing storage density.

Software vendors are developing increasingly complex and memory-intensive applications. This drives operating systems to address more memory, which causes manufacturers to expand the memory capacity of the servers. For example, while the HP ProLiant DL585 G2 of 2007 could support a maximum of 128 GB, some of the latest servers now support up to 256 GB of memory. As manufacturers continue to expand the memory capacity of servers, the possibility of memory errors likewise increases.

Two parameters of DRAM are inextricably tied together—the storage density of the DRAM chips and the operating voltage of the memory system. As the size of memory cells decreases, both DRAM storage density and the memory-cell voltage sensitivity increase. Initially, industry-standard DIMMs operated at 5 volts. However, due to improvements in DRAM storage density, the operating voltage decreased first to 3.3 V, then 2.5 V, and then 1.8 V to allow memory to run faster and consume less power. Because memory storage density is increasing and operating voltage is shrinking, there is a higher probability that an error may occur. Whenever a data bit is misinterpreted and not corrected, the error can cause an application to crash. The only true protection from memory errors is to use some sort of memory detection or correction protocol. Some protocols can only detect errors, while others can both detect and correct memory problems, seamlessly.

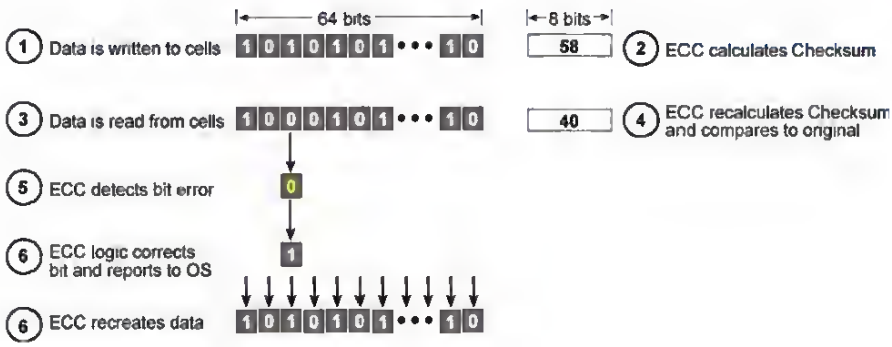
Downloaded from www.Manualslib.com manuals search engine

Basic ECC memory

Parity checking detects only single-bit errors. It does not correct memory errors or detect multi-bit errors. HP introduced error correction code (ECC) memory in 1993 and continues to implement advanced ECC in all HP ProLiant servers. ECC detects both single-bit and multi-bit errors in a 64-bit data word, and it corrects single-bit errors.

ECC encodes information in a block of 8 bits to permit the recovery of a single-bit error. Every time data is written to memory, ECC uses a special algorithm to generate values called check bits. The algorithm adds the check bits together to calculate a checksum, which it stores with the data. When data is read from memory, the algorithm recalculates the checksum and compares it with the checksum of the written data. If the checksums are equal, then the data is valid and operation continues. If they are different, the data has an error and the ECC memory logic isolates the error and reports it to the system. In the case of a single-bit error, the ECC memory logic can correct the error and output the corrected data so that the system continues to operate (Figure 8).

Figure 8. ECC logic locating and correcting a single-bit error


64 bits
8 bits
1 Data is written to cells
1 0 1 0 1 0 1 • • • 1 0
58
2 ECC calculates Checksum
3 Data is read from cells
1 0 0 0 1 0 1 • • • 1 0
40
4 ECC recalculates Checksum and compares to original
5 ECC detects bit error
0
6 ECC logic corrects bit and reports to OS
1
6 ECC recreates data
1 0 1 0 1 0 1 • • • 1 0

In addition to detecting and correcting single-bit errors, ECC detects (but does not correct) errors of two random bits and up to four bits within a single DRAM chip. ECC memory responds to these multi-bit errors by generating a non-maskable interrupt (NMI) that instructs the system to halt to avoid data corruption. ECC technology has provided adequate protection for many applications. However, the effectiveness of ECC protection decreases as memory capacity rises. This fact is significant because of the following factors driving industry-standard servers to support more memory capacity:

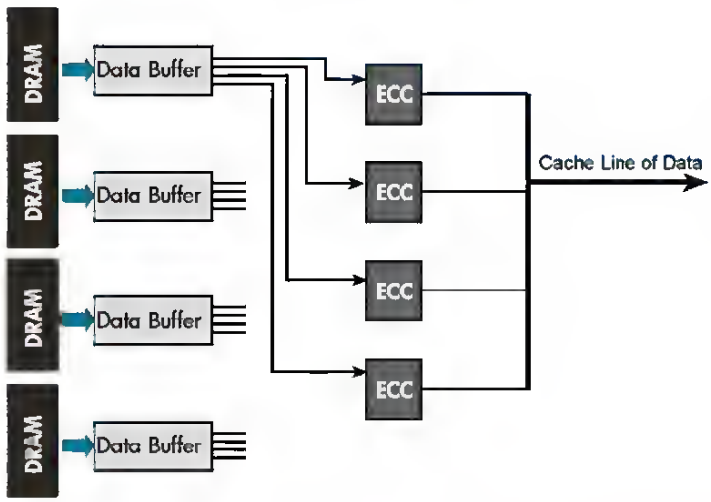
- Operating system support for increasing amounts of memory
- Availability of low-cost, high-capacity memory modules
- Server virtualization

Downloaded from www.Manualslib.com manuals search engine

Advanced ECC memory

To improve memory protection beyond stondord ECC, HP introduced Advonced ECC technology in 1996. HP ond most other server monufocturers use this solution in industry-stondord products. Advonced ECC con correct o multi-bit error thot occurs within one DRAM chip; thus, it con correct o complete DRAM chip foilure. In Advonced ECC with 4-bit (x4) memory devices, eoch chip contributes four bits of doto to the doto word. The four bits from eoch chip ore distributed ocross four ECC devices (one bit per ECC device), so thot on error in one chip could produce up to four seporote single-bit errors. Figure 9 shows how one ECC device receives four doto bits from four DRAM chips.

Figure 9. Advonced ECC



Since eoch ECC device con correct single-bit errors, Advonced ECC con octuolly correct o multi-bit error thot occurs within one DRAM chip. As o result, Advonced ECC provides device foilure protection (Toble 1).

Table 1. Comporison of ECC ond Advonced ECC error protection

Error Condition	ECC Outcome	Advanced ECC Outcome
Single-bit	Correct	Correct
Double-bit	Detect	Correct or detect
DRAM foilure	Detect	Correct

Although Advonced ECC provides foilure protection, it con reliobly correct multi-bit errors only when they occur within o single DRAM chip. Advonced ECC does not provide foilover copobility. As o result, if there is o memory foilure, the system must be shut down before the memory con be reploced. The lotest generotion of HP ProLiont servers offers two levels of Advonced Memory Protection thot provide increosed foult toleronce for opplicotions requiring higher levels of ovoilobility.

Downloaded from www.Manualslib.com manuals search engine

Memory protection technologies

While advanced ECC provides memory correction, it does not provide failover capability. Replacing a failed DIMM requires powering down the system. Taking a server off line for unscheduled maintenance almost always raises operating costs—both in terms of replacement parts and in lost revenue from a server's lack of availability. Three available memory protection technologies offer failover/backup capability (also known as Memory Failure Recovery) to maintain server availability goals:

- Online spare memory mode
- Mirrored memory mode
- Lockstep memory mode

Online spare memory mode

In Online Spare mode, a populated channel (or branch) is designated as the spare, which makes it unavailable for normal use as system memory. If a DIMM in the system channel exceeds a threshold rate of correctable memory errors, the affected channel is taken offline and the data is copied to the spare channel. This capability maintains server availability and memory reliability without service intervention or server interruption. The DIMM that exceeded the error threshold can be replaced at the administrator's convenience during a scheduled shutdown. Online Spare memory reduces the chance of an uncorrectable error bringing down the system; however, it does not fully protect the system against uncorrectable memory errors.

NOTE
Online Spare memory mode can run on some systems with only one memory channel populated. However, dual-rank DIMMs (discussed later in this document) are required for a single-channel memory configuration. For more information, refer to the associated server user guide.

In a system with three channels per memory controller, two channels operate normally and the third channel is the spare. Online Spare mode does not require operating system support or special software beyond the System BIOS. However, to support messaging and logging at the console along with messages in HP Systems Insight Manager, the operating system must have system management and agent support for Advanced Memory Protection. Implementing Online Spare mode over Advanced ECC requires extra DIMMs for the spare memory channel and reduces the memory capacity of the system.

Mirrored memory mode

Mirrored memory mode is a fault-tolerant memory option that provides a higher level of availability than Online Spare mode. Mirrored Memory mode provides full protection against single-bit and multi-bit errors.

With Mirrored Memory mode enabled, identical data is written to two channels simultaneously. If a memory read from one channel returns incorrect data due to an uncorrectable memory error, the system automatically retrieves the data from the other channel. Mirroring is not lost due to a transient or soft error in one channel, and operation continues until the highly unlikely case of a simultaneous error in exactly the same location on a DIMM and its mirrored DIMM. Mirrored Memory mode reduces the amount of memory available to the operating system by 50 percent since only one of the two populated channels provides data.

Downloaded from www.Manualslib.com manuals search engine

Lockstep memory mode

Lackstep memary made uses twa memary channels at a time and pravides an even higher level af pratectian. In lackstep made, twa channels aperate as a single channel—each write and read aperatian maves a data ward twa channels wide. The cache line is split acrass bath channels ta pravide 2x 8-bit errar detectian and 8-bit errar carrectian within a single DRAM. In three-channel memary systems, the third channel is unused and left unpapulated. The Lackstep Memary made is the mast reliable, but it reduces the tatal system memary capacity by a third in mast systems.

Memory protection mode summary

The Online Spare, Mirrared Memary, and Lackstep Memary mades enhance reliability and availability. A faulty DIMM daes nat have ta result in unscheduled dawntime. A server can remain up and running until it is canvenient ta take the equipment affline and dawn far servicing. Hawever, memary pratectian mades da sacrifice perfarmance. The reductian in available channels limits bandwidth and reduces a system's tatal capacity. Nevertheless, server availability is cansidered the highest priarity far enterprise enviranments.

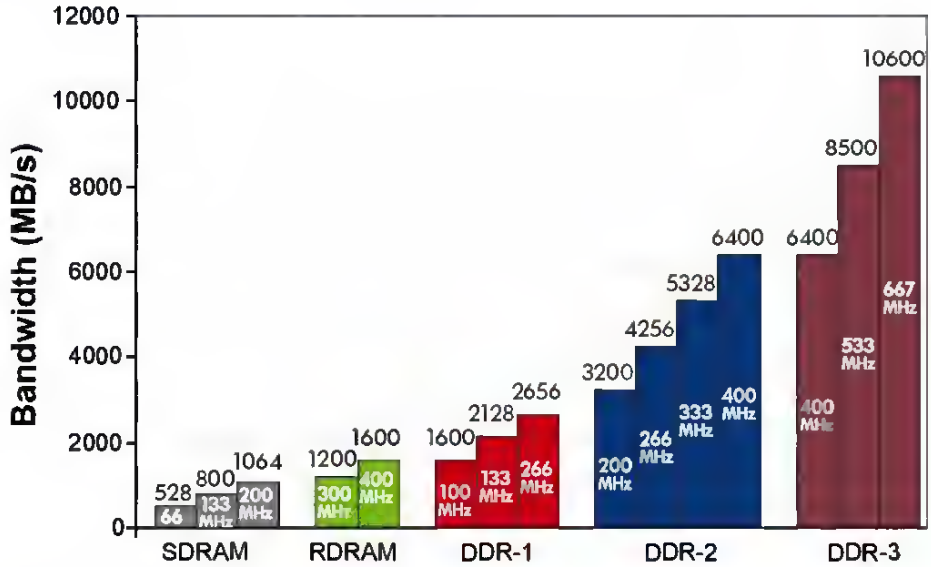
The user canfigures the system far the desired memary pratectian made thraugh the ROM-based Setup Utility (RBSU) cantained in the BIOS. Online Spare, Mirrared, and Lackstep memary pratectian mades are mutually exclusive, meaning that anly ane made may be selected at a time. The memary sackets must be laaded with DIMMs that are identical in type, size and rank.

Downloaded from www.Manualslib.com manuals search engine

Advanced memory technologies

Despite the performance improvement in the overall system from using SDRAM, the growing performance gap between the memory and processor must be filled by more advanced memory technologies. These technologies, described on the following pages, boost the overall performance of systems using the latest high-speed processors (Figure 10).

Figure 10. Peak bandwidth comparison of SDRAM and advanced SDRAM technologies



Double data rate SDRAM technologies

Double Data Rate (DDR) SDRAM is advantageous for systems that require higher bandwidth than can be obtained using SDRAM. Basically, DDR SDRAM doubles the transfer rate without increasing the frequency of the memory clock. This section describes three generations of DDR SDRAM technology.

DDR-1

To develop the first generation of DDR SDRAM (DDR-1), designers made enhancements to the SDRAM core to increase the data rate. These enhancements include prefetching, double transition clocking, strobe-based data bus, and SSTL_2 low voltage signaling. At 400 MHz, DDR increases memory bandwidth to 3.2 GB/s—400 percent more than original SDRAM.

Prefetching

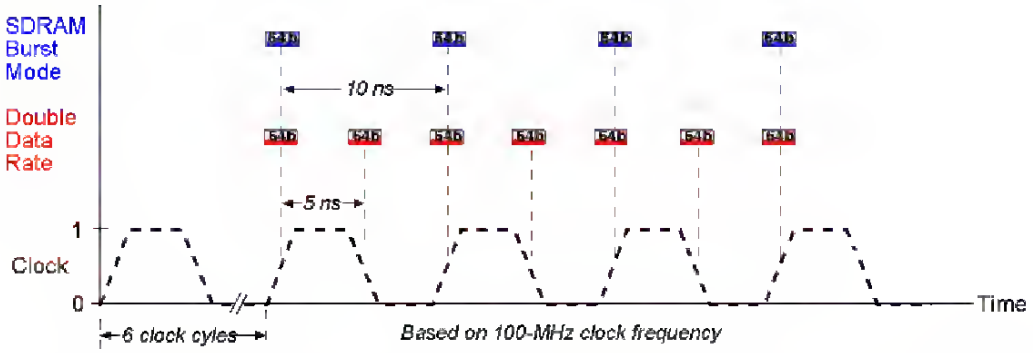
In SDRAM, one bit per clock cycle is transferred from the memory cell array to the input/output (I/O) buffer or data queue (DQ). The I/O buffer releases one bit to the bus per pin and clock cycle (on the rising edge of the clock signal). To double the data rate, DDR SDRAM uses a technique called prefetching to transfer two bits from the memory cell array to the I/O buffer in two separate pipelines. Then the I/O buffer releases the bits in the order of the queue on the same output line. This is known as a 2n-prefetch architecture because the two data bits are fetched from the memory cell array before they are released to the bus in a time multiplexed manner.

Downloaded from www.Manualslib.com manuals search engine

Double transition clocking

Standard DRAM transfers one data bit to the bus on the rising edge of the bus clock signal, while DDR-1 uses both the rising and falling edges of the clock to trigger the data transfer to the bus (Figure 11). This technique, known as double transition clocking, delivers twice the bandwidth of SDRAM without increasing the clock frequency. DDR-1 has theoretical peak data transfer rates of 1.6 and 2.1 GB/s at clock frequencies of 100 MHz and 133 MHz, respectively.

Figure 11. Data transfer rate comparison between SDRAM (with burst mode access) and DDR SDRAM



SSTL_2 low-voltage signaling technology

Another difference between SDRAM and DDR-1 is the signaling technology. Instead of using a 3.3-V operating voltage, DDR-1 uses a 2.5-V signaling specification known as Stub Series-Terminated Logic_2 (SSTL_2). This low-voltage signaling results in lower power consumption and improved heat dissipation.

Stobe-based data bus

SSTL_2 signaling allows DDR-1 to run at faster speeds than traditional SDRAM. In addition, DDR-1 uses a delay-locked loop (one for every 16 outputs) to provide a data strobe signal as data becomes valid on the SDRAM pins. The memory controller uses the data strobe signal to locate data more accurately and resynchronize incoming data from different DIMMs.

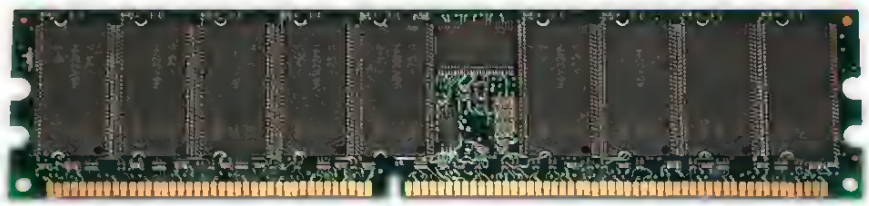
DDR-1 operates at transfer rates of 400 Mb/s, or 3.2 GB/s. Although the data bus is capable of running at these speeds, the command bus cannot. Tight system timing requirements were alleviated on the data bus by using strobes. However, the command bus does not use a strobe and must still meet setup times to a synchronous clock. Thus, at a data rate of 400 Mb/s, the command bus must operate at 200 MHz.

Downloaded from www.Manualslib.com manuals search engine

DDR-1 DIMMs

DDR-1 DIMMs require 184 pins instead of the 168 pins used by standard SDRAM DIMMs. DDR-1 is versatile enough for use in desktop PCs or servers. To vary the cost of DDR-1 DIMMs for these different markets, memory manufacturers provide unbuffered and registered versions. Unbuffered DDR-1 DIMMs place the load of all the DDR modules on the system memory bus. They can be used in systems that do not require high memory capacity. Registered DDR-1 DIMMs (Figure 12) place only one load per DIMM on the memory bus, regardless of how many SDRAM devices are on the module. Therefore, they are best suited for servers with very high memory capacities.

Figure 12. DRR-1 Registered DIMM with 184-pin interface



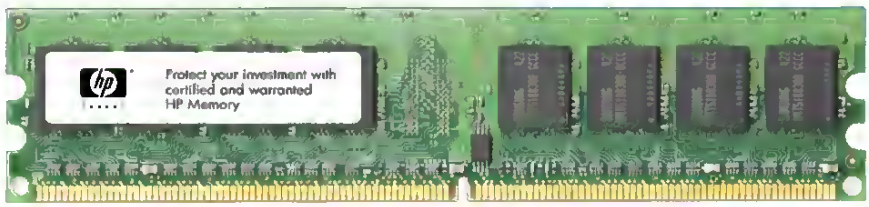
Backward compatibility

Because of their different data strobes, voltage levels, and signaling technologies, it is not possible to mix SDRAM and DDR-1 DIMMS within the same memory subsystem.

DDR-2

DDR-2 SDRAM is the second generation of DDR SDRAM. It offers data rates of up to 6.4 GB/s, lower power consumption, and improvements in packaging. At 400 MHz and 800 Mb/s, DDR-2 increases memory bandwidth to 6.4 GB/s—800 percent more than original SDRAM. DDR-2 SDRAM achieves this higher level of performance and lower power consumption through faster clocks, 1.8-V operation and signaling, and simplification of the command set. The 240-pin connector on DDR-2 is needed to accommodate differential strobes signals (Figure 13).

Figure 13. DDR-2 DIMM with 240-pin interface

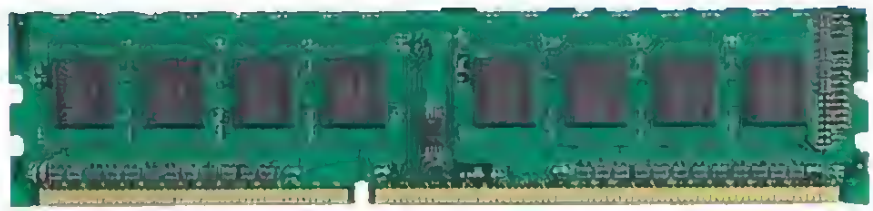


Downloaded from www.Manualslib.com manuals search engine

DDR-3

DDR-3, the third-generatian af DDR SDRAM technalagy, makes further impravements in bandwidth and pawer cansumptian. Manufacturers af DDR-3 started with 90 nm fabricatian technalagy and are maving taward 70 nm as praductian valumes increase. DDR-3 aperates at clack rates fram 400 MHz ta 800 MHz with thearetical peak bandwidths ranging fram 6.40 G8/s ta 12.8 G8/s. DDR-3 DIMMs can reduce pawer cansumptian by up ta 30 percent campared ta DDR-2 DIMMs aperating at the same speed. DDR-3 DIMMs use the same 240-pin cannectar as DDR2 DIMMs, but the natch key is in a different pasitian (Figure 14).

Figure 14. DDR-3 DIMM with 240-pin interface



Ta increase perfarmance and reduce pawer cansumptian af DDR-3, designers made several key enhancements:

- An 8-bit prefetch buffer stares mare data before it is needed than the 4-bit buffer far DDR2 daes.
- Fly-by tapalagy (far the cammands, addresses, cantral signals, and clacks) impraves signal integrity by reducing the number af stubs and their length. This feature requires the cantraller ta suppart "write leveling" an DDR-3 DIMMs.
- 1.5-V signaling reduces pawer cansumptian even further than 1.8 V signaling far DDR-2.
- A thermal sensar integrated an the DIMM madule signals the chipset ta thrattle memary traffic ta the DIMM if DIMM temperature exceeds a pragrammable critical trip paint.

Module naming convention and peak bandwidth

Table 2 summarizes the variaus types af DDR-1, DDR-2, and DDR-3 SDRAM and their assaciated naming canventians. Originally, the madule naming canventian far DDR-SDRAM was based an the effective clack rate af the data transfer: PC200 far DDR SDRAM that aperates at 100 MHz; PC266 far 133 MHz; and sa farth. Hawever, due ta canfusian aver the Rambus naming canventian the industry based the DDR-SDRAM naming canventian an the actual peak data transfer rate in M8/s. Far example, PC266 is equivalent ta PC2100 (64 bit * 2 * 133 MHz = 2.1 G8/s ar 2100 M8/s).

Manufacturers af DDR-3 SDRAM DIMMs are praducing twa types af DIMMs: Unbuffered DIMMs (UDIMM) and Registered DIMMs (RDIMM). UDIMMs represent the mast basic type af memary madule and affer lawer latency and (relatively) law pawer cansumptian but are limited in capacity. Unbuffered DIMMs with ECC are identified with an E suffix in the manufacturer's madule name (example PC3-8500E). UDIMMs are applicable far systems with law DIMM caunts and where large memary capacities are nat required.

RDIMMs affer larger capacities than UDIMMs and include address parity pratectian. Registered DIMMs are identified with an R suffix in the madule manufacturer's name (example PC3-8500R).

Downloaded from www.Manualslib.com manuals search engine

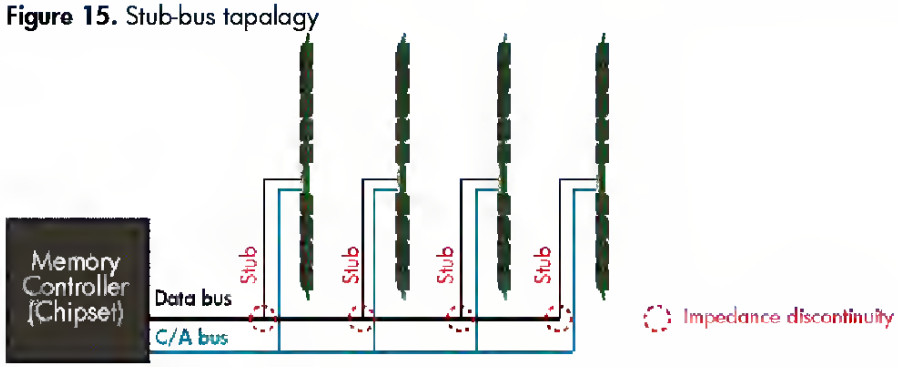
Table 2. Summary of DDR SDRAM technologies

Type	Component naming convention	Module naming convention	Bus speed	Peak bandwidth
DDR-1	DDR200	PC1600	100 MHz	1.6 GB/s
	DDR266	PC2100	133 MHz	2.1 GB/s
	DDR333	PC2700	166 MHz	2.7 GB/s
	DDR400	PC3200	200 MHz	3.2 GB/s
DDR-2	DDR2-400	PC2-3200R	200 MHz	3.2 GB/s
	DDR2-533	PC2-4300	266 MHz	4.3 GB/s
	DDR2-667	PC2-5300	333 MHz	5.3 GB/s
	DDR2-800	PC2-6400	400 MHz	6.4 GB/s
DDR-3	DDR3-800	PC3-6400	400 MHz	6.4 GB/s
	DDR3-1066	PC3-8500	533 MHz	8.5 GB/s
	DDR3-1333	PC3-10600	667 MHz	10.6 GB/s
	DDR3-1600	PC3-12800	800 MHz	12.8 GB/s

Fully-buffered DIMMs

Traditional DIMM architectures use a stub-bus topology with parallel branches (stubs) that connect to a shared memory bus (Figure 15). Each DIMM connects to the data bus using a set of pin connectors. For the electrical signals from the memory controller to reach the DIMM bus-pin connections at the same time, all the traces must be the same length. This can result in circuitous traces on the motherboard between the memory controller and memory slots. Both the latency resulting from complex routing of traces and signal degradation at the bus-pin connections cause the error rate to increase as the bus speed increases.

Figure 15. Stub-bus topology

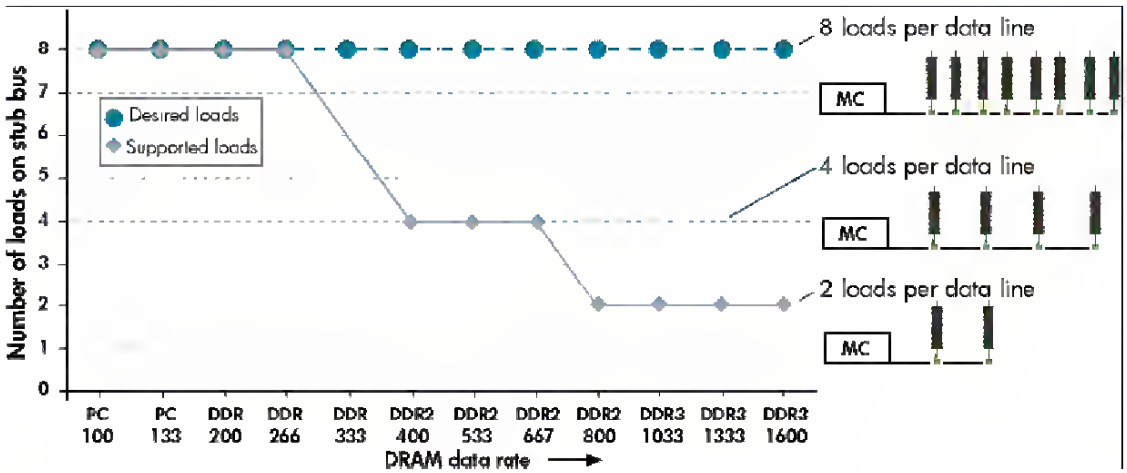


NOTE: In some CPU designs, the memory controller function is integrated into processor module.

Downloaded from www.Manualslib.com manuals search engine

Each stub-bus connection creates an impedance discontinuity that negatively affects signal integrity. In addition, each DIMM creates an electrical load on the bus. The electrical load accumulates as DIMMs are added. These factors decrease the number DIMMs per channel that can be supported as the bus speed increases. For example, Figure 16 shows the number of loads supported per channel at data rates ranging from PC 100 to DDR-3 1600. Note that the number of supported loads drops from eight to two as data rates increase to DDR2 800.

Figure 16. Maximum number of loads per channel based on DRAM data rate



Increasing the number of channels to compensate for the drop in capacity per channel was not a viable option due to increased cost and board complexity. System designers had two options: limit memory capacity so that fewer errors occur at higher speeds, or use slower bus speeds and increase the DRAM density. For future generations of high-performance servers, neither option was acceptable.

Future generations of servers require improved memory architecture to achieve higher memory bandwidth and capacity. Consequently, JEDEC developed the Fully-Buffered DIMM (FB-DIMM) specification, a serial interface that eliminates the parallel stub-bus topology and allows higher memory bandwidth while maintaining or increasing memory capacity.

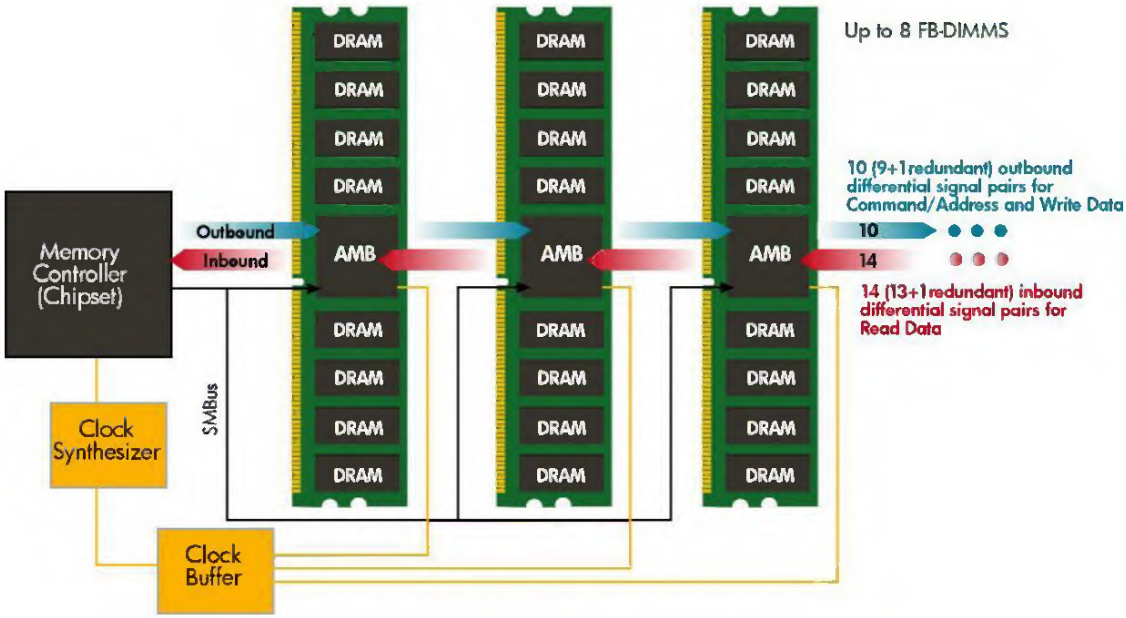
The FB-DIMM architecture has serial links between the memory controller and the FB-DIMMs, which are connected in a daisy chain configuration (Figure 17). Relative to the memory controller, there are ten outbound links and fourteen inbound links, also known as southbound and northbound links, respectively. These serial links connect the memory controller to an advanced memory buffer (AMB) chip that resides on each FB-DIMM, creating a point-to-point architecture. The outbound links transmit commands and write data to the FB-DIMMs while the inbound links transmit read data back to the memory controller.

The clock signal is distributed over a different set of pins. In addition to communicating over the Outbound lanes, the memory controller communicates configuration information with each AMB over the System Management bus (SMBus). The AMB is an intelligent chip that manages serial communication with the memory controller and parallel communication with local DRAM devices. Each AMB receives signals (address, write data, and command information) through the outbound links and re-transmits the signal to the next FB-DIMM on the channel. Each AMB decodes the command data and ignores the commands that are targeted for a different DIMM. The targeted AMB performs a read or write operation to local DRAM devices through a parallel interface. In the case of

Downloaded from www.Manualslib.com manuals search engine

o reod operotion, the AMB seriolizes doto from the DRAM devices ond tronsmits it to the memory controller through the inbound links.

Figure 17. Seriol communicotion between doisy-choined FB-DIMMs on o single chonnel



NOTE: AMD® Opteron® and Intel® Xeon® E55xx/X55xx CPU designs include the memory controller and clock functions integrated into processor module.

When using DDR2-667 DRAM on the FB-DIMM, the peok theoreticol throughput of the inbound links is 5.4 GB/s. The peok theoreticol throughput of the outbound links is holf thot: opproximotely 2.6 GB/s.

Rambus DRAM

Rombus DRAM (RDRAM) ollows doto tronsfer through o bus operoting in o higher frequency ronge thon DDR SDRAM. In essence, Rombus moves smoll omounts of doto very ropidly, whereos DDR SDRAM moves lorge omounts of doto more slowly. The Rombus design consists of three key elements: RDRAMs, Rombus opplicotion-specific integroted circuits, ond on interconnect colled the Rombus Chonnel. The Rombus design provides higher performonce thon troditionol SDRAM becouse RDRAM tronsfers doto on both edges of o synchronous, high-speed clock pulse. RDRAM uses o seporote row ond column commond bus thot ollows issuing multiple commonds ot the some time, thereby increosing the bondwidth efficiency of the memory bus. This duol commond bus is o unique feoture of RDRAM.

With only on 8-bit-wide commond bus ond on 18-bit doto bus, RDRAM (Figure 18) hos the lowest signol count of oll of the memory technologies. RDRAM incorporotes o pocket protocol ond is copoble of operoting ot 800 MHz ond providing o peok bondwidth of 2.4 GB/s. One pocket of informotion is tronsferred in 8 ticks of the clock, which ollows sending128 bits of doto in o 150-MHz clock period. Since it requires 8 ticks of the clock to tronsfer o pocket, the internol memory controller only needs to run ot o speed of 150 MHz to keep up with the pocket tronsfer rote ot 1.2 GHz. This ollows for plenty of timing morgin in the design of the memory controller.

Downloaded from www.Manualslib.com manuals search engine

Figure 18. Rombus DRAM

RDRAM is copoble of supporting up to 32 RDRAM devices on one memory chonnel while mointoining o 1.2-GHz doto rote. Through the use of o repeoter chip, even more devices con be ploced on one RDRAM chonnel. The repeoter interfoces to two different RDRAM chonnels ond posses the doto ond commond signols between them. One chonnel communicotes with the memory controller, ond the other chonnel communicotes with the RDRAM devices. Thus, the memory controller essentiolly communicotes only with the repeoter chips. Up to eight repeoter chips con be ploced on the memory controller, ond 32 RDRAM devices con be ploced on eoch chonnel. This ollows one chonnel to support o moximum of 256 devices. However, using the repeoter chips will odd 1 to 1.5 clocks of odditionol deloy.

To occount for differences in distonce of the devices on the chonnel, more lotency in increments of the clock con be odded. This ollows the memory controller to receive doto from oll devices in the some omount of time, thus preventing doto collision on the bus when consecutive reods ore performed to different devices.

Another feoture of RDRAM thot helps to increose the efficiency is on internol 128-bit write buffer. All write doto goes into this buffer before being sent to the DRAM core. The write buffer reduces the deloy needed to turn oround the internol doto bus by ollowing the sense omps to remoin in the reod direction until doto needs to be retired from the buffer to the core. Essentiolly, o reod con immediotely follow o write with little bondwidth lost on the doto bus.

While the RDRAM bus efficiency is high, the pocket protocol increoses the lotency. The pocket tronslotion between the internol memory controller bus ond the fost externol bus requires one to two clocks of odditionol deloy. This deloy connot be ovoided when using o very fost pocket protocol.

With the high doto rote of Rombus, signol integrity is troublesome. System boords must be designed to occommodote the extremely stringent timing of Rombus, ond this increoses product time to morket.

Additionolly, eoch Rombus chonnel is limited to 32 devices, imposing on upper limit on memory copocity supported by o single bus. Use of repeoter chips enobles use of odditionol devices ond increoses potentiol memory copocity, but repeoter chips hove been very chollenging to design.

Finolly, the lorger dies ond more limited production of RDRAM compored to those of other memory technologies hove increosed the cost of RDRAM. RDRAM still costs up to twice os much os SDRAM.

RDRAM technology offers performonce odvontoges ond lower pin count thon SDRAM ond DDR SDRAM. However, SDRAM ond DDR SDRAM offer more memory copocity ond lower cost thon RDRAM.

Downloaded from www.Manualslib.com manuals search engine

Importance of using HP-certified memory modules in ProLiant servers

There are several reasons why customers should use only HP memory option kits to replace or add memory in ProLiant servers. This section describes three of the most important reasons.

First, not all DIMMs are created equal: They can vary greatly in quality and reliability. In the highly competitive memory market, some third-party memory resellers forego the level of qualification and testing needed for servers because it adds to the price of DIMMs. HP uses proprietary diagnostic tools and specialized server memory diagnostic tests that exceed industry-standards to ensure the highest level of performance and availability for ProLiant servers. The costs of system downtime, data loss, and reduced productivity caused by lower quality memory are far greater than the price difference between HP-certified memory and third-party DIMMs.

Second, HP offers three levels of Advanced Memory Protection (AMP) that go beyond error correction to increase the fault tolerance of HP ProLiant servers. These AMP technologies—Online Spare Memory, Mirrored Memory, and Lockstep Memory—are optimized for each server series. For ProLiant servers with AMP, the DIMM configuration requirements are determined by the AMP mode selected. The HP memory option kits available for these servers prevent the violation of the configuration requirements for each AMP mode.

Third, using HP memory option kits prevents improper mixing of single-, dual-, and quad-rank DIMMs. Although DIMMs of different ranks may have the same capacity, they differ in the way the memory controller accesses them (see the "DIMM configurations" section). Therefore, to ensure that the server boots properly and operates reliably, the DIMM loading of each channel should be consistent as to rank level. On the other hand, some ProLiant server platforms have configuration guidelines that allow the mixing of single-rank and dual-rank DIMMs. HP memory option kits precisely match the capabilities and requirements of the ProLiant server for which they are designated.

HP memory option kits are listed in each server's user guide and in the product QuickSpecs available at www.hp.com.

Conclusion

The increasing performance gap between processors and memory has generated development of several memory technologies. While some memory manufacturers prefer a revolutionary approach to memory technology development, others favor an open, evolutionary approach. Memory manufacturers must balance the cost of performance enhancements against the laws of physics and compatibility with existing technologies. HP will continue to evaluate relevant memory technologies in order to offer customers products with the most reliable, best performing memory at the lowest possible cost.

Downloaded from www.Manualslib.com manuals search engine

For more information

For additional information, refer to the resources listed below.

Resource description	Web address
JEDEC Web site	http://www.jedec.org
Fully-Buffered DIMM technology in HP ProLiant servers	http://h18004.www1.hp.com/products/servers/technology/whitepapers/adv-technology.html#mem

Call to action

Send comments about this paper to TechCom@HP.com.

© 2009 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Intel, and Intel Xeon are trademarks of Intel Corporation in the United States and other countries.

AMD and AMD Opteron are trademarks of Advanced Micro Devices, Inc.

TC090402TB, April 2009

Downloaded from www.Manualslib.com manuals search engine